DOCUMENT RESUME

ED 195 562 TM 810 002

AUTHOR Plake, Barbara Sterrett: And Others
TITLE Comparing Profile Variability Across Independent Groups.
PUB DATE Sep 80
NOTE 12p.: Paper presented at the Annual Meeting of the American Psychological Association (88th, Montreal, Canada, September 1-5, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Diagnosis: Educational Testing: *Individual Testing: *Mathematical Models: *Measurement Techniques: Psychological Testing: *Scores: *Statistical Analysis: Statistical Significance: Test Interpretation
IDENTIFIERS *Profile Variability: *Test Scatter: Wechsler Intelligence Scale for Children (Revised)

ABSTRACT

The difficulties in comparing profile variability (a measure of test scatter) are briefly discussed and the limitations of current techniques pointed out. Test scatter is defined as individual variation in test scores between or within various psychological and educational tests. Currently, no statistical technique for the comparison of profile variability across groups is available. Profile variability is explained in detail and a statistical technique is developed to compare profile variability across independent groups, i.e., the variance of each subscale or subtest within an instrument. Although development of the statistic is accomplished with applications to the Wechsler Intelligence Scale for Children (Revised) in mind, it is generalizable to any set of test scores derived for an individual. Theoretical and practical applications of the new technique to diagnosis, and description of abnormal populations are discussed. (Author/RL)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED195562

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

# Comparing Profile Variability Across Independent Groups

Barbara Sterrett Plake, Cecil R. Reynolds, and Terry B. Gutkin

University of Nebraska

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

B. Plake

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the annual meeting of the

American Psychological Association, Montreal, 1980

TM 810002

## Abstract

The difficulties in comparing profile variability (a measure of test scatter) are briefly discussed and the limitations of current techniques pointed out. A powerful technique is then developed that takes into account the variance of each subscale or subtest within an instrument. Theoretical and practical applications of the new technique to diagnosis and description of abnormal populations are discussed.

Comparing Profile Variability Across Independent Groups

Clinical, school, and other psychologists frequently involved in assessment have long believed that a large degree of test scatter (individual variation in test scores between or within various psychological and educational tests) characterizes certain types of psychological disorders (Waugh & Bush, 1971; Zimmerman & Woo-Sam, 1973). The concept of abnormal scatter has been especially prevalent in various definitions of learning disability (Chalfant & Scheffelin, 1969). Recently however, Kaufman (1976 a,b) has reported data demonstrating that normal children exhibit much scatter in their test profiles, calling into question the utility of scatter indices in the differential diagnosis of mental disorders. Test scatter is typically determined by the range (highest minus lowest score) or by the number of test scores deviating significantly from a mean of all tests administered to the individual. Profile variability is another measure of test scatter, similar to the range, but more powerful; profile variability encompasses data from all tests or subtests administered rather than being limited to the two most extreme scores, as is the range statistic.

Several recent studies have sought to compare test scatter across samples of normal and exceptional children by comparing the mean range of subtest scores of these groups on the WISC-R (Anderson, Kaufman, & Kaufman, 1976; Gutkin, in press). While statistically significant differences in range of subtest scores have been shown to occur between

normal children and children diagnosed as learning disabled, emotionally disturbed, brain injured, and mildly retarded, the small actual size of the differences makes them of questionable diagnostic utility. Profile variability is a potentially more powerful index since it retains information from all tests (or subtests) administered rather than relying on the highest and lowest tests only. Profile variability should be a more stable index of scatter as well since it relies on a constant set of scores for each individual. The range of scores for each individual in a group may well be determined by two different scores for each person. Other measures, such as the variance of total scores, are also less sensitive to profile scatter than a measure of profile variance. The index of abnormality ($A_d$) has also been proposed as a means of evaluating test scatter (Piotrowski, 1978), however this technique has also been shown to be inappropriate (Reynolds, in press).

Currently, no statistical technique for the comparison of profile variability across groups is available. Profile variability is explained in detail below and a statistical technique developed to compare profile variability across independent groups. Although development of the statistic is accomplished with applications to the WISC-R in mind, it is generalizable to any set of test scores derived for an individual.

## Development of the Statistic

Assume the subscale of interest is composed of K subtests, each of which have separate scores reported. The profile variability for each subject in the population could be estimated by:

$$S_i^2 = \frac{\sum_{j=1}^{K} (X_{ij} - \bar{X}_{i.})^2}{K-1}$$

where:

$X_{ij}$ is the score for subject i on subtest j.

$\bar{X}_{i.}$ is the mean for subject i of the K scores from the subscale.

K is the number of subtests in the subscale (or tests in the battery).

Under the assumptions of normally distributed scores and independent random sampling, $\frac{(K-1)\, S_i^2}{\sigma_i^2}$ will have a $\chi^2$ distribution with K-1 degrees of freedom. When the scores are reported in a normalized fashion, the normal assumption is met. If the subtests that make-up the subscale are nearly independent, the second assumption will be satisfied. However, since the variance covariance matrix of the tests comprising a subscale are likely to display dependence between tests, there is a possibility of violations of this assumption. However, the statistic will be robust in this regard.

For a group of N subjects, each of whom have taken the K tests that comprise the subscale,

$$\sum_{i=1}^{N} \sum_{j=1}^{K} \frac{(X_{ij} - \bar{X}_{i.})^2}{\sigma_i^2}$$

is a $\chi^2$ with N (K-1) degrees of freedom due to the additive property of independent $\chi^2$.

Simplifying we have:

$$\sum_{i=1}^{N} \sum_{j=1}^{K} \frac{(X_{ij} - \bar{X}_{i.})^2}{\sigma_i^2} = (K-1) \sum_{i=1}^{N} \frac{S_i^2}{\sigma_i^2}$$

If we assume that all the subjects within the population have equal profile variances, then we can further simplify:

$$\frac{K-1}{\sigma^2} \sum_{i=1}^{N} S_i^2$$

For two independent groups of subjects, two independent $\chi^2$ variables can be determined as follows:

Group I:

$$\frac{K-1}{\sigma_i^2} \sum_{i=1}^{N_1} S_i^2 \sim \chi^2_{N_2 (K-1)}$$

Group II:

$$\frac{K-i}{\sigma_2^2} \sum_{i=1}^{N_2} S_2^2 \sim \chi^2_{N_2 (K-1)}$$

The two $\chi^2$ variables are independent since they are based on independent groups. The ratio of the two independent $\chi^2$ variables, each divided by their respective degrees of freedom, will form an F ratio:

$$\underline{F}_{N_1 (K-1),\ N_2 (K-1)} = \frac{\frac{\frac{K-1}{\sigma_1^2} \sum_{i=1}^{N_1} S_{i1}^2}{N_1 (K-1)}}{\frac{\frac{K-1}{\sigma_2^2} \sum_{i=1}^{N_2} S_{i2}^2}{N_2 (K-1)}}$$

Under the following set of assumptions, this $\underline{F}$ ratio will follow an $\underline{F}$ distribution with $N_1 (K-1)$ and $N_2 (K-1)$ degrees of freedom.

Assumptions:

1) Independent random sampling.

2) Profile scores are normally distributed.

3) Homogeneity of variance within groups.

The first assumption, independent random sampling, is the basic assumption for all parametric statistical procedures. The research sampling design will guard against violations of this basic assumption.

The second assumption, normal distribution of profile scores holds for the total population of test takers if the profile scores are standardized. For a specific group of subjects, a check on the distribution of scores may be warranted, especially since the $\underline{F}$ statistic is sensitive to violations of the assumption of normal distribution.

The third assumption, homogeneity of within group profile variance must be met before the test statistic exactly follows an $\underline{F}$ distribution. It has been shown (Box, 1954) that with nearly equal sample sizes, the $\underline{F}$ statistic is robust to violations of this type.

Under the null hypothesis, $H_o : \sigma_1^2 = \sigma_2^2$, the $F$ ratio reduces to:

$$F_{N_1 (K-1),\, N_2 (K-1)} = \frac{\dfrac{\dfrac{K-1}{\sigma_1^2} \sum_{i=1}^{N_1} S_{i1}^2}{N_1 (K-1)}}{\dfrac{\dfrac{K-1}{\sigma_2^2} \sum_{i=1}^{N_2} S_{i2}^2}{N_2 (K-1)}} = \frac{\dfrac{\sum_{i=1}^{N_1} S_{i1}^2}{N_1}}{\dfrac{\sum_{i=1}^{N_2} S_{i2}^2}{N_2}} = \frac{\bar{S}_1^2}{\bar{S}_2^2}$$

## Applicability of Statistic

It is often desirable to investigate the variability of two populations of subjects across a subset of tests. The hypothesis of interest is represented as:

$H_o: \sigma_1^2 = \sigma_2^2$ where $\sigma_1^2$ is the profile variance for population 1 and $\sigma_2^2$ is the profile variance for population 2.

To test this hypothesis, two equal sized groups would be randomly sampled independently from the two populations of interest. For each group, the sum of the profile variances divided by N would be calculated. Their ratio would be the test statistic used to test the hypothesis of interest. Specifically, the test statistic is:

$$F_{N_1 (K-1),\, N_2 (K-1)} = \frac{\bar{S}_1^2}{\bar{S}_2^2}$$

## Discussion

For comparing test scatter between normal and pathological or other populations, the profile range has typically been employed. The profile

variance ($S_j^2$) is a better index of profile variability since it utilizes all subtest scores rather than the extreme, and is based on a constant set of observations for each individual within all groups of concern. Employing the test statistic described herein, it is possible to statistically assess the equality of profile variances across several groups, allowing researchers to ascertain statistical significance of observed differences in profile variability. Being a highly sensitive measure, profile variability measured through the technique described herein may also have substantial diagnostic utility. The use of the profile variability approach may help to clarify the massive number of contradictory findings reported in previous research (Matarrazzo, 1972; Sattler, 1974) which utilized less stable statistical techniques.

The measure of profile variability is also appropriate for use in the comparison of behavior ratings when within-subject variability is of interest. As one part of a validation study for a classroom observation code, Abikoff, Gittelman-Klein, and Klein (1977) compared the within-subjects variability of observational ratings across five categories of behavior. In their study, correlated *t*-tests were employed to compare the mean range of ratings across the five categories. Only two of the five categories were then allowed to influence the outcome. The comparison of profile variability proposed herein allows input from all five categories and is more powerful, efficient, and precise in answering questions of within-subjects variability when more than two scales, subtests, or categories are employed.

## References

Abikoff, H., Gittelman-Klein, R., & Klein, D. F. Validation of a classroom observation code for hyperactive children. Journal of Consulting and Clinical Psychology, 1977, 45, 772-783.

Anderson, M., Kaufman, A. S., & Kaufman, N. L. Use of the WISC-R with a learning disabled population: Some diagnostic implications. Psychology in the Schools, 1976, 13, 381-386.

Box, G.E.T. Some theorems on quadratic forms applied in the study of analysis of variance problems. Annals of Mathematical Statistics, 1954, 25, 290-302.

Chalfant, J. C., & Scheffelin, M. A. (eds.) Central processing dysfunctions in children: A review of research, phase 3. NI-NDs Monograph No. 9. Washington, D.C. : U.S. Government Printing Office, 1969.

Gutkin, T. B. WISC-R scatter indices: Useful information for differential diagnosis? Journal of School Psychology, in press.

Kaufman, A. . A new approach to the interpretation of test scatter on the WISC-R. Journal of Learning Disabilities, 1976, 9, 160-167.

Kaufman, A. S. Verbal-Performance IQ discrepancies on the WISC-R. Journal of Consulting and Clinical Psychology, 1976, 44, 739-744.

Matarazzo, J. D. Wechsler's measurement and appraisal of adult intelligence. Baltimore: Williams & Wilkins Co., 1972.

Piotrowski, R. S. Abnormality of subtest score differences on the WISC-R. Journal of Consulting and Clinical Psychology, 1978, 46, 569-570.

Reynolds, C. R. Interpreting the index of abnormality ($A_d$) when the distribution of score differences is known: Comment on Piotrowski. Journal of Consulting and Clinical Psychology, in press.

Sattler, J. M. Assessment of children's intelligence. Philadelphia: W. B. Saunders, 1974.

Waugh, W. J., & Bush, K. W. Diagnosing learning disorders. Columbus, Ohio: Charles Merrill, 1971.

Zimmerman, I. L., & Woo-Sam, J. M. Clinical interpretation of the Wechsler Adult Intelligence Scale. New York: Grune and Stratton, 1973.